SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES

The effect of Encryption on Audio Steganography

أثر التشفير في إخفاء المعلومات في الصوت

February 2015

SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE S TUDIES

The effect of Encryption on Audio Steganography

أثر التشفير في إخفاء المعلومات في الصوت

A Thesis Submitted in Partial Fulfillment of the Requirements of Master Degree in Computer Science

BY: SUPERVISOR

AFAF YOUSIF AHMED DR. TALAAT WAHBY

February 2015

الآيـــــــــــــــــــــــة

قال تعالى(وَمَا أُوتِيتُمْ مِنْ الْعِلْمِ إِلاَّ قَلِيلاً)

صدق الله العظيم

DETICATION

To my family

To my teachers

To my best friends

ACKNOWLEDGEMENTS

All praise to Allah first and last for helping me to complete this thesis.

I would like to express my sincere thanks to my supervisor for providing me precious advices, guidance and suggestions.

Also I would like to express greater thanks to my family who were always supporting and encouraging me.

Special respect and thanks to my teacher my uncle Abdurrahman Ahmed, Mustafa Mohammed, Suhaip Ali and all members group for their support and help.

ABSTRACT

Cryptography and Steganography are the two popular methods available to provide security. Recently a combination between them becomes the most successful ways to maintain confidentiality of data but with impact on quality.

LSB is a common method to hide data but it is easy for attacker to extract the hidden message.

This thesis concentrates and focuses on the impact of encryption (both symmetric and asymmetric) on the quality of hiding messages in audio files. An LSB method has been developed to avoid the lack of quality.

It is clear from the evaluation of the results that both symmetric and asymmetric encryption affect the quality of hiding data by 0.02 and 6.98 %, respectively, compared with hiding plaintexts. It is also found that the proposed method, AFT_LSB, has increased and improved the quality of hiding messages compared with standards LSB.

المستخـــلص

مؤخراً أصبح الإخفاء والتشفيرمن أكثر الطرق شيوعاً للمحافظة على سرية البيانات،وقد أصبح الدمج بينهما من أنجح الوسائل للمحافظه على سرية البيانات المخفاه.

يركز هذا البحث على دراسة اثر التشفير بنوعيه المتماثل وغير المتماثل على جودة إخفاء البيانات في ملف صوتي بإستخدام خوارزمية البت الأقل أهميه (LSB) كما يركز على تطوير خوارزمية البت الأقل أهميه (Standard LSB) للتقليل من عيوبها إضافة الى قياس أثر إخفاء رسالة محفوظه بصيغ مختلفة (txt.docx,pdf) على جودة الصوت .

بعد دراسة النتائج إتضح أن جودة الصوت تاثرت بنسبة 0.2%عند تشفير الرساله بخوارزمية AES وبنسبة 6.98% عند تشفيرها بخوارزمية RSA مقارنة مع إخفاء الرسالة من غير تشفير وأن الخوارزمية المقترحه (AFT_LSB) زادت جودة الإخفاء بنسبة 38.51%مقارنة مع Standards LSB .

Table Of Content

List Of Figure

List of Table

INTRODUCTION

1.1 Introduction

Recently the exchange of data over the Internet became indispensable. As a result, it has became important to use hiding and encryption mechanics to maintain the privacy and confidentiality of data as it passed across the network.

Information hiding is the technology that is used to embed the secret information into a cover data in a way that keeps the secret information invisible. There are many techniques that can be used for hiding data. One of these techniques is Steganography.

Steganography, coming from the Greek words stegos, meaning roof or covered and graphic, which means writing, is the art and science of hiding the fact that communication is taking place. Using Steganography, you can embed a secret message inside a piece of unsuspicious information and send it without anyone knowing the existence of the secret message.

Steganography can be a solution which makes it possible to send news and information without being censored and without the fear of the messages being intercepted and traced back to you. While sending messages can be useful, it is also possible to simply use Steganography to store information on a location. For example, several information like your private banking information and some military secrets, can be stored in a cover source. When you are required to unhide the secret information in your cover source, you can easily reveal your banking data and it will be impossible to prove the existence of the military secrets inside [**1**].

Steganography can be used a large amount of data formats in the digital world of today. The most popular data formats are .bmp, .doc, .gif, .jpeg, mp3, .txt and .wav. Steganography technologies are a very important part of the future of Internet security and privacy on open systems such as Internet [**2**] .

Fundamentally, audio Steganography is the art and science of hiding digital data such as text messages, documents, and binary files into audio files such as WAV, MP3, and

RM files. The output audio file is called the carrier file and is the only intermediate to be sent to the receiver. Characteristically, anyone who is taping the communication wire would not notice anything suspicious being transmitted, except for a normal audio file this property of Steganography is called imperceptibility and it refers to the fact that no one apart from the original sender and the intended receiver can suspect the presence of secret data into the carrier file being communicated [**3**] .

Figure 1.1 below illustrates the basic idea of any Steganography process

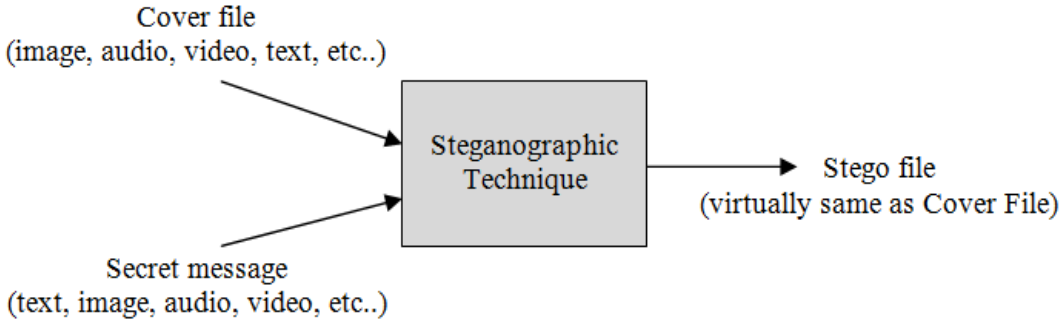


Figure 0.1 fundamental scheme of Steganography process

Steganography and cryptography are closely related. Cryptography scrambles messages so they cannot be understood. Steganography, on the other hand, hides the messages so that there is no knowledge of the existence of data in the first place.

Steganography and cryptography are both ways to protect information from unwanted parties. But Steganography differs from cryptography in the sense that where cryptography focuses on keeping the contents of a message secret, Steganography focuses on keeping the existence of a message secret

Steganography is not the same as cryptography. Data hiding techniques have been widely used for transmitting secret message for long time. Ensuring data security is a big challenge for computer users. Business men, professionals, and home users all have some important data that they want to secure from others. Even though both methods provide security, to add multiple layers of security it is always a good

practice to use Cryptography and Steganography together. The combination of these two methods will enhance the security of the data embedded.

A pictorial representation of the combined concept of cryptography and Steganography is explained in figure 1.2 [**4**] .

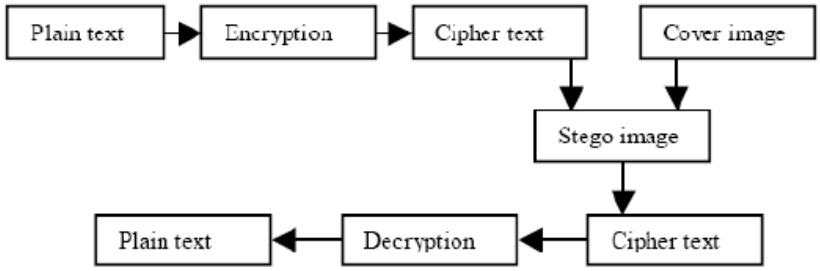


Figure 0.2 combination of Steganography and cryptography

1.2 Problem Statement

Hiding data in an audio file using the LSB algorithm depends on replacing LSB value in audio with a message consecutively. This process makes it **easy for attackers to extract** the hidden message. Once an attacker has the ability to extract hidden message that means the failure of Steganography process.

The strength of Steganography can be amplified by combining it with cryptography to provide a very acceptable amount of secrecy. However, the question of **what is the impact of encryption on the quality of concealment is an important question to ask.**

1.3 Thesis Objectives

The objectives of this thesis are:

- Enhancing LSB algorithm to hide message in audio to reduce noise.
- Studying the impact of encryption on the hiding.

- Trying to hide message in wav file without affecting the quality of audio.
- Implementing Standard LSB to hide data in audio.
- Extracting the message safely from the stegoFile.
- Comparing the quality of hiding message using our enhanced algorithm with the one using Standard LSB Steganography algorithms.
- Trying to find the appropriate cryptography algorithm to combine it with the new LSB without affecting the quality of hiding.

1.4 Thesis methodology and tools

After reading deeply and analyzing previous studies in the field, I have designed a model that combines cryptography with Standard LSB Steganography. In addition, ~~to~~ I designed a new LSB method to hide message in wave file so as to avoid the weakness of LSB algorithm.

Moreover, several experiments have been conducted to study the effect of encryption on the quality of concealment and compare the quality of my algorithm with LSB Standard using MSE and PSNR metrics.

In this thesis, Java has been selected for implementation. This because it is a high-level language and it is probably the most secure programming language to date. Furthermore, the Java Class Library provides a number of APIs related to security, such as standard cryptographic algorithms, authentication, and secure communication protocols.

1.5 Scope of Thesis

The scope is illustrated in Figure 1.3:

- A new technique is implemented to hide encrypted file (.docx, .text, .pdf) in audio file (.wav)
- AES symmetric algorithm is used to encrypt file.
- RSA asymmetric algorithm is used to encrypt file.
- LSB algorithm is used to hide data in audio File.

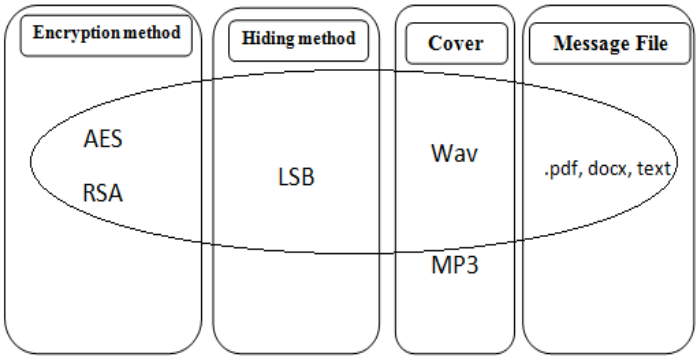


Figure 0.3 Scope of Thesis

1.6 Thesis Structure

This thesis contains five chapters. Literature review and related work are presented in chapter two. AFT_LSB method is the main topic of Chapter three. Experiments' result and Discussions are discussed in chapter four. Finally, Recommendation and Future works are discussed in the last chapter.

.

LITERATURE REVIEW AND RELATED WORKS

2 Overview

The chapter begins with the definition and some background of Steganography and cryptography next; it presents the related works of LSB audio Steganography.

2.1 Introduction

Information security is a major issue of concern while exchanging a data in an open network, as internet is not only a single network it is worldwide collection of loosely network. It is not restricted by any geographical, national or international boundaries; it means anyone can access it from any part of the world. Although it is very useful for various purposes but there is a risk associated with security of the information which is transfer through the internet. Anyone can hack the information and then make misuse from that or corrupt it or we can say that anyone can destroy the information if it is not fully secured or protected.

The security of information passed over an open channel has become a fundamental issue and therefore, the confidentiality and data integrity are required to protect against unauthorized access and use [**4**]. Steganography and Cryptography both plays a very important role in information security one hides the existence of the message and the other distorts the message itself. [**5**]

2.2 Steganography

Generally speaking, Steganography brings science to the art of hiding information. Steganography, coming from the Greek words stegos, which mean roof or covered and graphic which means writing, is the art and science of hiding the fact that communications taking place. Using Steganography, you can embed a secret message inside apiece of unsuspicious information and send it without anyone knowing of the existence of the secret message. [**1**]

In Steganography systems the following terms are used:

Cover Media: The cover media is the medium in which message is embedded. StegoFile: The media through which the data is hidden.

Secret data: The data to be hidden or extract.

Stego key or simply key (k): This is additional embedded secret data which may be needed in the information hiding process. In particular, this key (or a related one) is typically needed to extract [**6**].

2.2.1 Steganography Cover media

Almost all digital multimedia files - text, image, audio, video and protocol - can be used as cover mediums for Steganography to hide secret data .but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object's use and display. The redundant bits of an object are those bits that can be altered without the alteration being detected easily. Image and audio files especially comply with this requirement. .Figure 2.1 shows the four main categories of file formats that can be used for Steganography

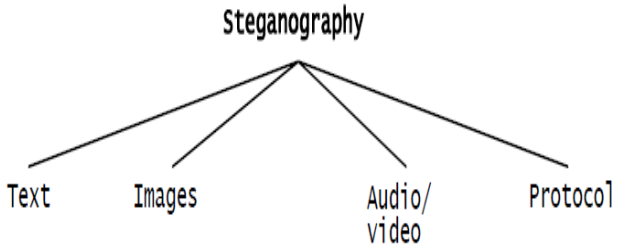


Figure 2.1 Categories of Steganography

2.2.1.1 Text Steganography

Hiding information in text is the most important method of Steganography. The method was to hide a secret message in every n th letter of every word of a text

message. After booming of Internet and different type of digital file formats it has decreased in importance. Text stenography using digital files is not used very often because the text files have a very small amount of redundant data.

2.2.1.2 Image Steganography

Images are used as the popular cover objects for Steganography. A message is embedded in a digital image through an embedding algorithm, using the secret key. The resulting Stego image is send to the receiver. On the other side, it is processed by the extraction algorithm using the same key. During the transmission of stego image unauthenticated persons can only notice the transmission of an image but can't guess the existence of the hidden message.

2.2.1.3 Audio Steganography

Audio stenography is masking, which exploits the properties of the human ear to hide information unnoticeably. An audible, sound can be inaudible in the presence of another louder audible sound .This property allows to select the channel in which to hide information

2.2.1.4 Protocol Steganography

The term protocol Steganography is to embedding information within network protocols such as TCP/IP. We hide information in the header of a TCP/IP packet in some fields that can be either optional or are never used. [**7**]

In my Thesis will be concentrated on audio Steganography

2.2.2 Audio Steganography

When secret data is embedded into digital sound, the technique is known as audio Steganography. This method embeds the secret message in WAV, AU and MP3 sound files [**8**].

The secret message is concealed into the audio media by slightly changing the binary sequence of the audio file. In order to hide secret information successfully, a range of techniques for inserting information into digital audio have been introduced.

2.2.2.1 Steganographic techniques

There are many steganographic techniques for hiding secret data or messages in audio in a way that the modifications made to the audio file are perceptually indiscernible. One of the common techniques LSB coding.

❖ Least Significant Bit (LSB) Coding

Basically, the computer was created due to binary numbers, known as two numbers, namely 0 and 1. Both of these numbers are often referred to as bits. Then, these bits will continue to form a composite sequential and binary structure into a set of information. Set of information is composed of 8-bit or often referred to as 1 byte.

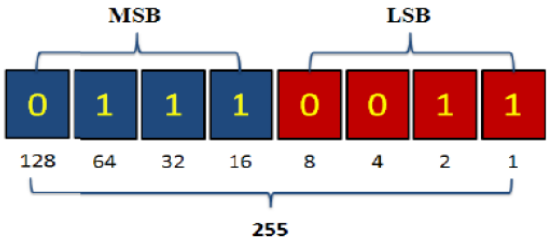


Figure 2.2 bits representation

Binary information bits are classifications based on the sequence and its influence in the byte bits. These bits are divided into 2 groups, the Most Significant Bit (MSB) and the LSB, as shown in figure 2.2 Most Significant Bit is representation of 4-bits which have a major influence on a range of information, means drastic changes that would occur if these bits are modified. While the Least Significant Bit is a 4-bit representation of the least influential when the bits are modified and will not be a drastic change, so the possibility of human prejudice against LSB bits are modified very little. Thus, the right, the bits are smaller effect on the integrity of the data contained. Therefore, the 4-bit last modified and became the sticking a Steganography digital information. [**9**]

Least significant bit coding is the easiest and simplest method to hide secret data in a digital audio media. By replacing the least significant bit of each sample words with a bit of the secret data, LSB coding permits a big size of secret data to be embedded. [**6**]

❖ Standard LSB Method

Least significant bit method is the simplest way to embed information in a digital audio file. By substituting the least significant bit of each sampling point with a binary message, LSB coding allows for a large amount of data to be encoded. [**10**].According to the number of replaced bits standard LSB divided into four types

a) One LSB

One least significant bits of a sample are replaced with one message bits. as shown in figure 2.3

Figure 2.3 One LSB Example

b) Two LSB

Two least significant bits of a sample are replaced with two message bits. This increases the amount of data that can be encoded but also increases the amount of resulting noise in the audio file as well. As shown in figure 2.4

Figure 2.4 Two LSB Example

c) Three LSB

Three least significant bits of a sample are replaced with three message bits. This increases the amount of data and noise more than one and two LSB. As shown in figure 2.5

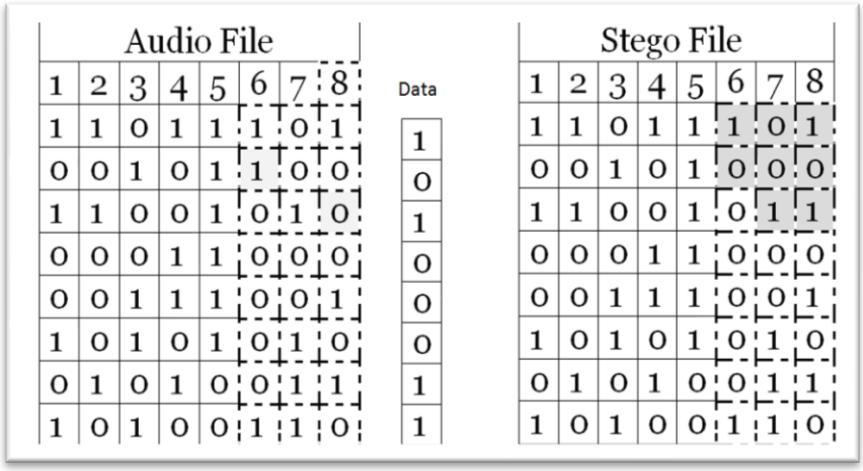


Figure 2.5 Three LSB Example

d) Four LSB

Four least significant bits of a sample are replaced with four message bits. This increases the amount of data and noise more than others LSB methods.

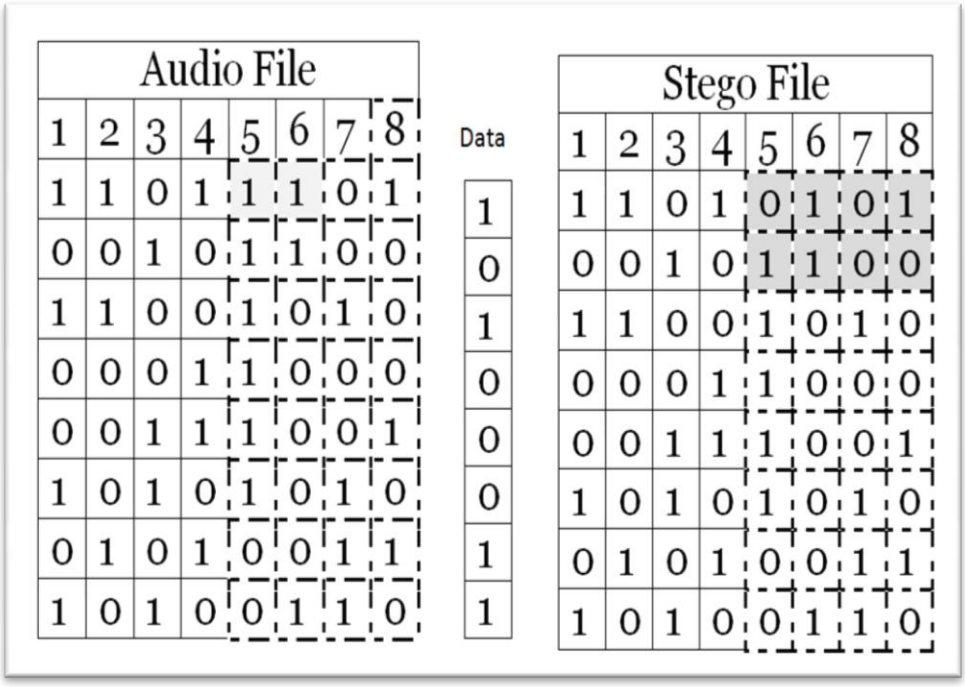


Figure 2.6 Four LSB Example

Advantage of LSB method

It is the simplest way to embed information in a digital audio file. It allows large amount of data to be concealed within an audio file. This method is more widely used as modifications to LSBs usually not create audible changes to the sounds.

Disadvantage of LSB method:

It has considerably low robustness against attacks [**11**].

2.3 Cryptography

Cryptography is art and science of keeping messages secure. it is the practice and study of techniques for secure communication in the presence of third parties. Cryptography protects information by transforming it into unreadable format. Only those who possess a secret key can decipher the cipher text into plain text.

2.3.1 Cryptography Terminology

In Cryptography systems following terms are used:

- Plaintext is message or data which are in their normal, readable form.
- Encryption: Encoding the contents of the message in such a way that hides its contents from outsiders.
- Cipher text: encrypted plaintext
- Decryption: The process of retrieving the plaintext back from the cipher text.
- Key: Encryption and decryption usually make use of a key, and the coding method is such that decryption can be performed only by knowing the proper key [**4**].

2.3.2 Types of Cryptography

There are two main categories of cryptography depending on the type of security keys used to encrypt/decrypt the data. These two categories are:

2.3.2.1 Symmetric-key cryptography

It is also called as single key cryptography. It uses a single key. In this encryption process the receiver and the sender has to agree upon a single secret (shared) key. Symmetric-key cryptography refers to encryption methods in which both the sender and receiver share the same key. In symmetric-key cryptography, the same key is used by both parties. The sender uses this key and an encryption algorithm to encrypt data; the receiver uses the same key and the corresponding decryption algorithm to decrypt the data. This was the only kind of encryption publicly known until June 1976.

Symmetric key ciphers are implemented as either block cipher or stream cipher. A block cipher ciphers input in blocks of plaintext as opposed to individual characters, the input form used by a stream cipher.

Stream ciphers, in contrast to the 'block' type, create an arbitrarily long stream of key material, which is combined with the plaintext bit-by-bit or character-by-character. In a stream cipher, the output stream is created based on a hidden internal state which changes as the cipher operates. That internal state is initially set up using the secret key material [**12**].

2.3.2.2 AES: (Advanced Encryption Standard)

AES also known as the Rijndael (pronounced as Rain Doll) algorithm, is a symmetric block cipher that can encrypt data blocks of 128 bits using symmetric keys 128, 192, or 256. AES was introduced to replace the DES. Brute force attack is the only effective attack known against this algorithm.

2.3.2.3 Asymmetric Encryption

It is also called as Public-key cryptography, where different keys are used for encryption and decryption. In asymmetric or public-key cryptography, there are two keys: a private key and a public key are used. The private key is kept by the receiver and public key is announced to the public.

2.3.2.4 RSA

RSA is an algorithm for public-key cryptography that is based on the presumed difficulty of factoring large integers. A user of RSA creates and then publishes the product of two large prime numbers, along with an auxiliary value, as their public key. The prime factors must be kept secret. Anyone can use the public key to encrypt a message.

RSA is an internet encryption and verification scheme and is the most commonly used algorithm. The algorithm engrosses multiplying two big prime numbers and by means

of additional operations derives a set of two numbers in which one set comprises the public key and other set comprises the private key. Both the public and the private keys are desired for encryption and decryption purposes but only the holder of private key desires to recognize it[**14**].

Table 2.1 Differences between Steganography and Cryptography

Steganography	Cryptography
Hides a message within another message and looks like a normal graphic, video, or sound file.	The message is encrypted; looks like a meaningless jumble of characters.
Collection of graphic images, video files, or sound files on a disk may not look suspicious.	Collection of random characters on a disk may look suspicious.
A smart eavesdropper can detect something suspicious from a sudden change of message format (i.e., text to graphic images).	A smart eavesdropper can detect a secret communication from a message that has been cryptographically encoded

2.4 Related works:

In this section we focus on the studies related with my problem. The authors are mainly concerned with the security of the embedded message in audio using LSB.

In [**16**], S.S. Divya, M. Ram Mohan Reddy proposed two novel approaches of LSBs of audio samples for data hiding. These methods check the MSBs of the samples, and then number of LSBs for data hiding is decided. These methods utilize upto 7 LSBs for embedding data. Results show that both these methods improve capacity of data hiding of cover audio by 35% to 70% as compared to the standard LSB algorithm with 4LSBs used for data embedding. And using encryption and decryption

techniques performing cryptography. So for this RSA algorithm used. This method analyzed in terms of PSNR, incr_cap (Increased Capacity) and MSE [**16**]

In [**17**], Padmashree G, Venugopala P S used the 4^{th} and 5^{th} Bits LSB method for embedding the message into the audio file. The quality of the audio file after encoding remains unaffected. RSA was also used to ensure greater security. In all the cases(Same Audio File With Varying Text Content Sizes ,Different Audio Files Of Different Time Durations With Same Text Content, Different Categories Of Audio File With Same Text Content), SNR and PSNR are calculated and The results show that the size of the audio file remains same even after embedding the secret message [**17**].

In [**18**], Sumod Tom Philip, SumayaNazar, Ashams Mathew &Niya Joseph proposed method to combination of cryptography using AES and Steganography is used here for security. To increase the security level , the key is hashed using SHA-1.the data encrypted using advanced encryption algorithm will be hidden into a multimedia image, audio and video file according to the user's choice. Steganography is implemented by means of Least Significant Bit insertion technique [**18**].

In [**19**], Kriti Saroha, Pradeep Kumar Singh proposed a new steganographic method for embedding an image in an Audio. Proposed method is found to be better than the 3LSB and 4LSB methods [**19**].

Table 2.2 related work summarization

Paper no	Features and Strength	Notes
16	Used MSB value to determine the LSB and RSA was used to increase security Increased the capacity As compared to standard LSB PSNR and MSE test	-
17	2LSB was used and Combine it with Cryptography(RSA) Different case was evaluated by PSNR.	Result not compared with Standard LSB
18	Hide text in an image, audio or video file. AES was used to increase security	test cases PSNR and MSE test
19	Use MSB bits to hide data in audio Difficult to recover message from attack PSNR and MSE test.	low capacity One level of security

PROPOSED METHOD

3 Introduction

This chapter explains the methodology applied in this thesis. It describes the overall stages of the experiment framework used.

3.1 Proposed Framework

In this thesis we will build a model to study the effect of encryption on the quality of concealment. Here we use an English message to conceal it into an audio file. In encoding process below the message is send to hidden process as plain text and as cipher text to generate StegoFile .The extraction process is an inverse of the encoding process to obtain the original message. As shown in Figure 3.1.

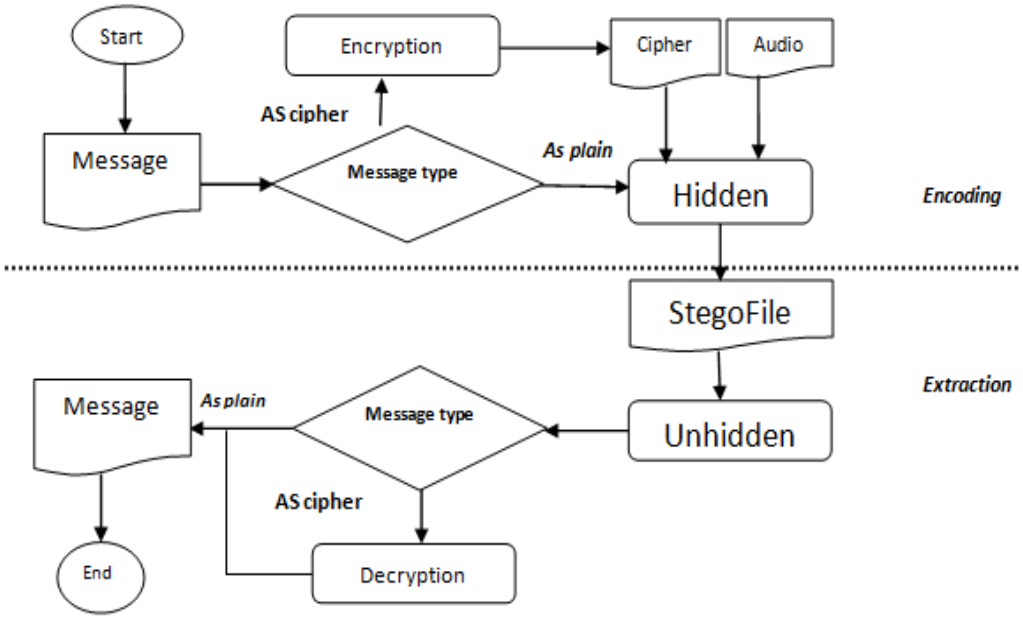


Figure 3.1 Proposed Framework

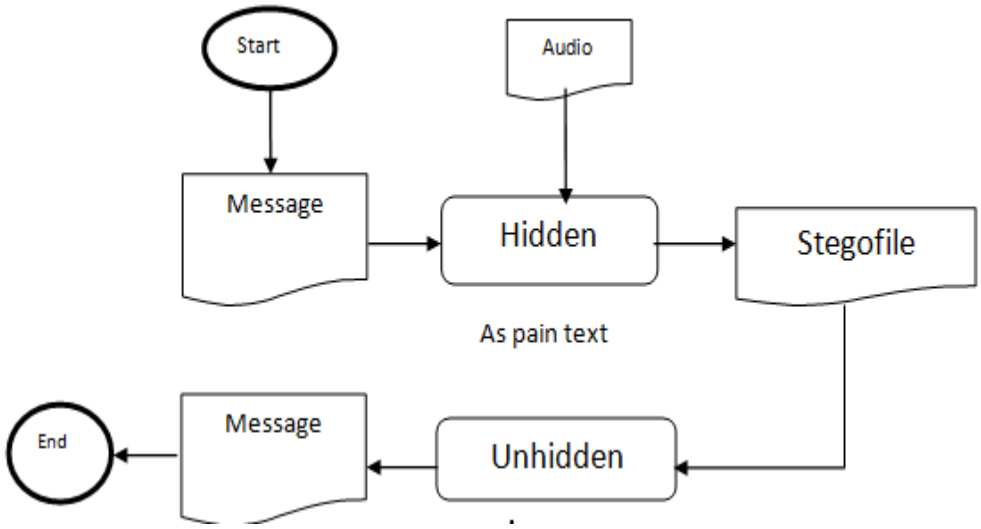


Figure 3.2 Plaintext Model

In the model 3.2 above the message (Plain text) is hidden directly in audio file.

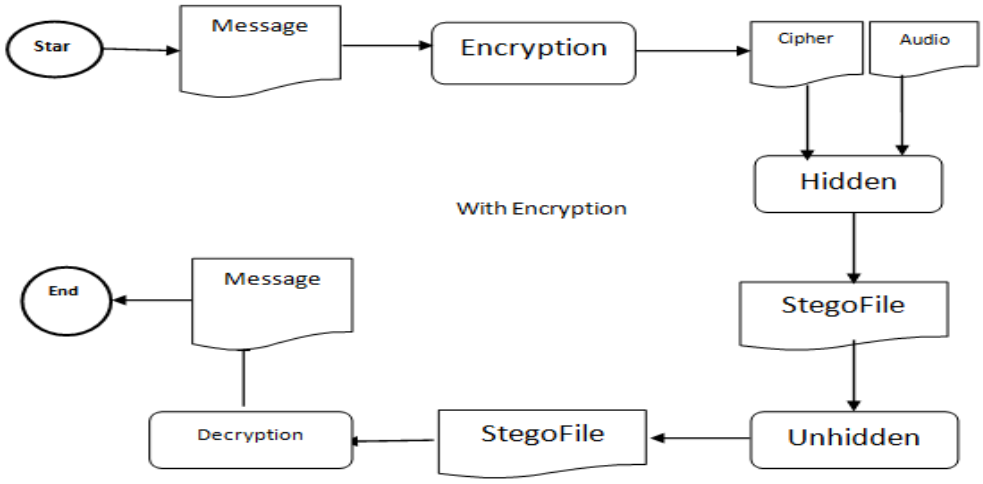


Figure 3.3 with encryption model

In model 3.3 above the message file is encrypted firstly using AES, and RSA algorithms. Then the outputs of (AES cipher and RSA cipher) are hidden in audio file.

The hidden process is done using StandardsLSB that discussed in chapter 2 section 2.2.2.2.2.and AFT_LSB will be discuss in this chapter.

3.2 AFT_LSB

AFT_LSB is my proposed method to reduce LSB weakness we used variable LSBs for embedding secret data in audio.

3.2.1 Stages of AFT_ LSB Method

AFT_LSB use two stages to hide message in audio. Firstly determine the number of message bits to be hidden in any byte. Secondly determine the number of jumping byte using MSBs.

3.2.1.1 Number of message bits

Check first and second bits (MSB value) for any bytes of audio to determine what number of message bits to be hides in it. By this way we can use variable Standard LSB to hide one message. As it explain in table 3.1 below. Example below illustrates how to determine the number of message bits.

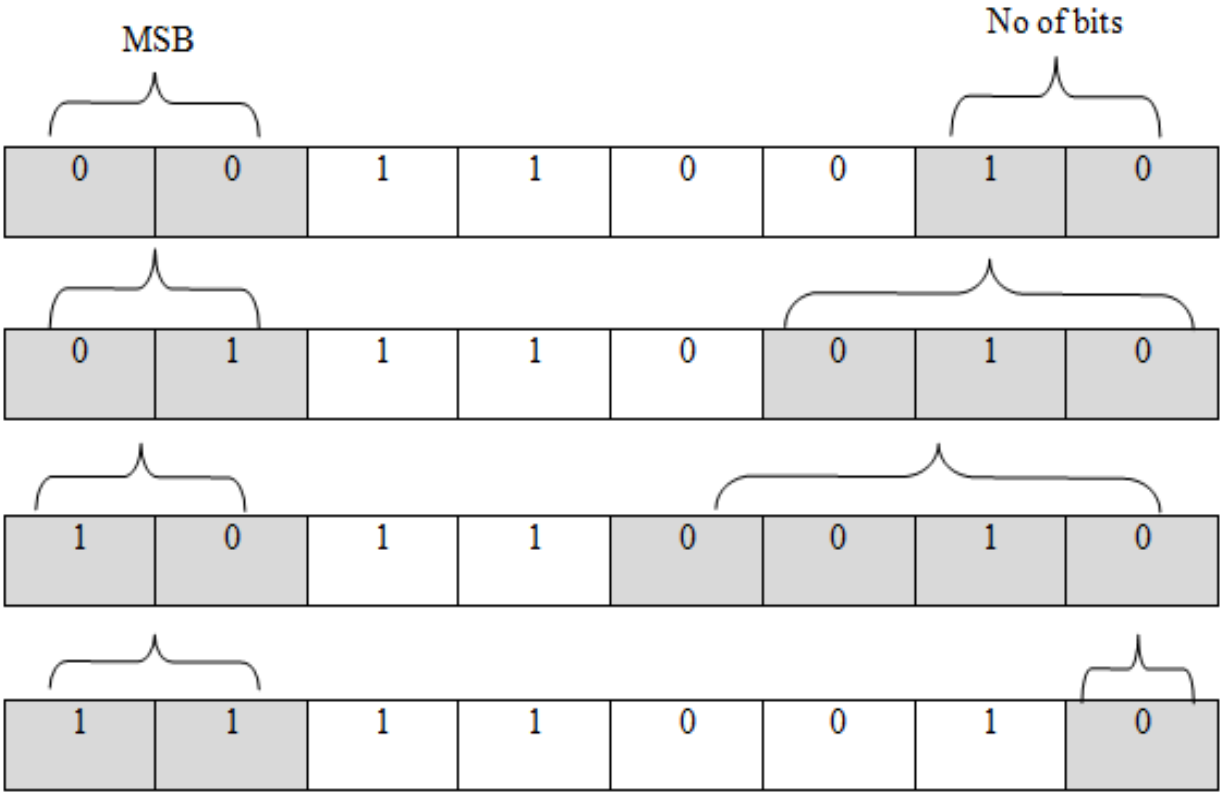


Figure 3.4 AFT_LSB no of bits

3.2.1.2 Jumping Techniques

To reduce noise and increase difficulty of extraction from attacker we must not store message consecutively by using jumping techniques.

According to the MSB value of the storage byte of audio decide what number of byte to be written in StegoFile without any modification. The jumping byte was generated according to the equation below:

$$\text{Jump}=\text{Length of message mod } ((no*2)+1) \quad \textbf{Equation 3-1}$$

If the result of the equation above is odd value the jumping bytes equals the same value of the equation. And if the result of the equation is even value jumping bytes equals the value of equation below:

$$\text{Jumping Byte}=\text{jump}+ ((no*2)+1) \quad \textbf{Equation 3-2}$$

Table 3.1 AFT table

No	MSB(2 bit)	No of LSBs bit
1	00	2
2	01	3
3	10	4
4	11	1

Example below illustrates how the message is encoded in audio using the AFT_LSB method. Firstly the secret information and the audio file are converted into bit stream.

Secondly the least significant bit of the audio file is replaced by the bit stream of secret information depending on the MSB value. The resulting file after embedding secret information is called Stego-file.

Table 3.2 AFT_LSB example

Message=8865 byte	1	1	0	1	0	0	1	1	0	1	1	0	0	0	0	1	0	1	1	0	0	0	1	0
No of byte	audio File																Stego File							
1	1	0	1	1	0	1	0	1									1	0	1	1	1	1	0	1
									Jump3 byte															
5	1	0	1	0	0	1	1	1									1	0	1	1	0	0	1	1
									Jump 3 byte															
9	0	0	1	1	1	1	1	1									1	1	1	1	1	1	0	1
									Jump3 byte															
13	1	1	0	0	0	0	0	1									0	0	0	0	0	0	0	1
									Jump9 byte															
23	0	0	1	1	0	1	1	1									1	1	1	1	1	1	0	0
									Jump3 byte															
27	0	1	0	0	0	0	0	0									0	0	0	0	0	0	1	0
									Jump5byte															
33	1	1	0	0	0	0	0	1									0	0	0	0	0	0	0	0
									Jump9 byte															
43	0	1	0	0	0	0	0	0									0	0	0	0	0	1	1	0
									Jump5 byte															
49	0	0	1	1	1	1	1	1									1	1	1	1	1	1	0	0
									Jump3byte															
53	0	0	1	1	1	1	1	1									1	1	1	1	1	1	1	0

To extract message from StegoFile Firstly convert Stego file to bit stream. Secondly extract least significant bit of the audio file depending on MSB value.

3.2.2 Steps of Data Embedding

The Steps of Data Embedding explain in Figure 3.2 are:

1) Choose encryption algorithm.

2) Encrypt text file

3) Calculate number of jump bytes

4) Read the header of audio file and write it in stegoFile

5) Read audio byte and convert it to bits

6) Read text byte and convert it to bits

7) Check bit1 and bit2 in audio byte

- If it is “00” then embed two bits of text and write a new byte in stegoFile write jumping bytes in stegoFile
- If it is “01” then embed three bits of text and write a new byte in stegoFile Write jumping bytes in stegoFile
- If it is “10” then embed four bits of text and write new byte in stegoFile
- If it is “11” then embedding one bits of text and write new byte in stegoFile Write jumping bytes in stegoFile

8) Repeat steps (5, 6, 7) until text file is ended. Figure 3.1 above illustrates Steps for Data Embedding.

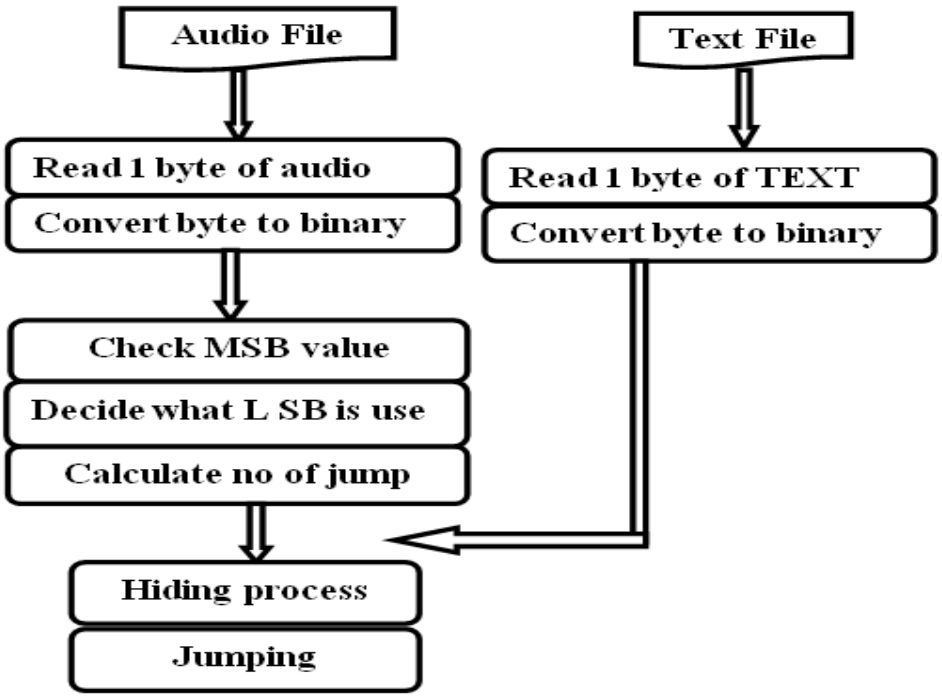


Figure 3.5Steps of Data Embedding

3.2.3 Steps of Data Retrieval

The retrieval steps illustrated in Figure 3.2 below are:

1) Calculate number of jump bytes

2) Read the stegoFile.

3) Ignore the header of audio file

4) Read audio byte and convert it to bits

5) Check the first bit and the second bit

- If it is “00” then extract two bit and ignore jump bytes
- If it is “01” then three bits and ignore jump bytes
- If it is “10” then embedding four bits and ignore jump bytes

- If it is “11” then embedding one bits and ignore jump bytes

6) Repeat steps 4,5 until number of hidden byte is ended

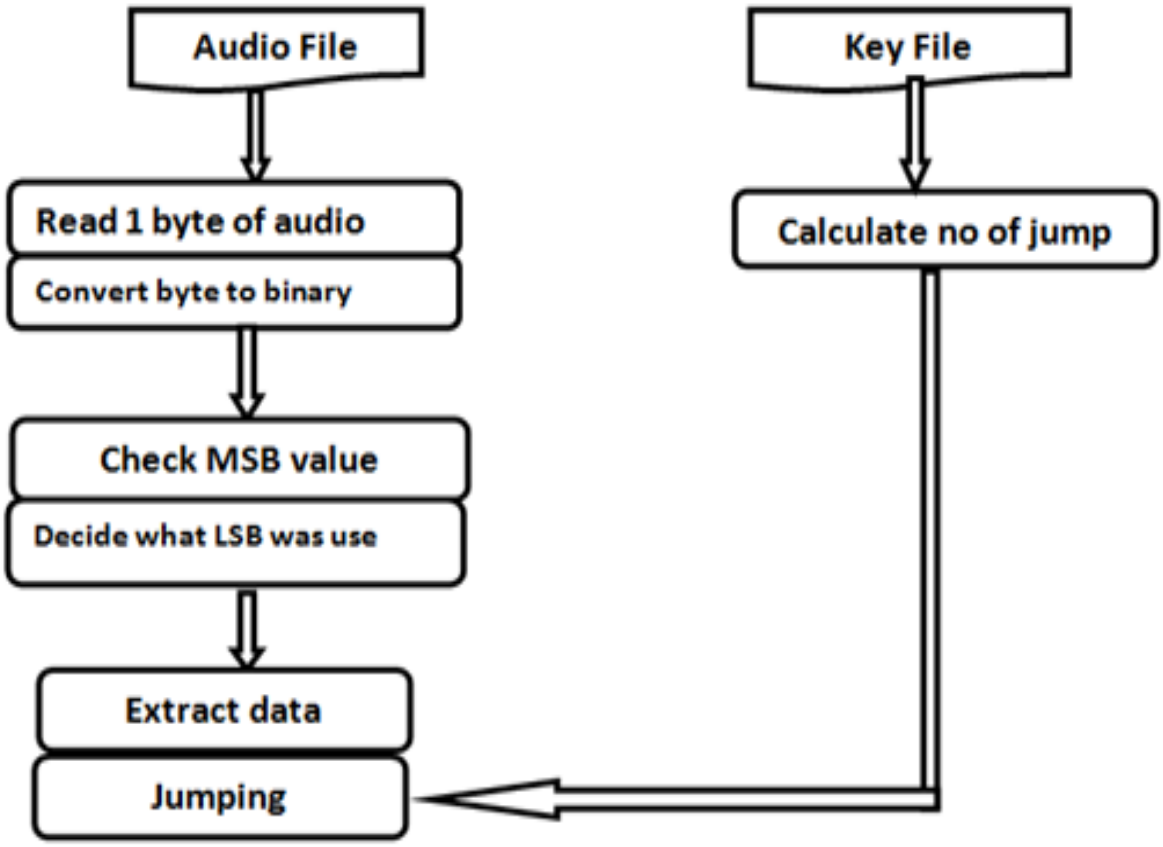


Figure 3.5 Steps for Data Retrieval

EXPERIMENTS RESLUTS

4 Introduction

This chapter discusses the results of the experiments which conducted to evaluate quality of StegoFile. Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) were used to for evaluation.

4.1 Mean Square Error

MSE can be measure is by compare two signals by providing a quantitative score that describes the degree of similarity or, conversely, the level of error/distortion between them. Usually, it is assumed that one of the signals is a pristine original, while the other is distorted or contaminated by errors [**21**].

$$\mathbf{MSE}=\sum_{\mathbf{n=0}}^{\mathbf{n}}[\mathbf{x(n)}-\mathbf{y(n)}]^{\mathbf{2}}$$

Here x (n) represents cover audio file and y (n) represents StegoFile.

4.2 Peak Signal-to-Noise Ratio

PSNR is an expression for the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. The PSNR is usually expressed in terms of the logarithmic decibel scale [**20**].

$$\mathbf{PSNR}=\mathbf{10}log_{\mathbf{10}}\frac{(\frac{B}{2})-\mathbf{1}}{MSE}$$

4.3 Experiments Result and discussion

In all experiments we use same audio file (45sec sample rate 16 bits) with different text file (plain text, AES cipher text, RSA cipher text) with varying text content sizes. We divided my experiments into two groups

1. Impact of Size message on PSNR value

To study impact of encryption and message size conducted three types of experiments with varying message sizes.

a) Message without encryption(plain text)

Table 4.1 Impact message size on PSNR value

method	1 LSB		2 LSB		3LSB		4LSB		AFT_LSB	
size	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
1	123.583	0.000000	118.122	0.000000	112.814	0.000000	107.823	0.000001	158.666	0.000000
2	120.534	0.000000	115.269	0.000000	109.708	0.000001	104.744	0.000002	155.869	0.000000
3	118.712	0.000000	113.703	0.000000	108.137	0.000001	102.792	0.000003	153.915	0.000000
8	114.184	0.000000	110.050	0.000001	104.377	0.000002	98.687	0.000009	149.022	0.000000

From the table 4.1 above it is clear that the PSNR value decrease when the message length and number of LSB bits increases

b) Message Encrypted by AES

Table 4.2 Impact Result of Size in AES of the PSNR value

method	1 LSB		2 LSB		3LSB		4LSB		AFT_LSB	
size	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
1	123.476	0.000000	118.180	0.000000	112.616	0.000000	107.450	0.000001	159.126	0.000000
2	120.499	0.000000	115.250	0.000000	109.769	0.000001	104.666	0.000002	155.809	0.000000
3	118.738	0.000000	113.452	0.000000	107.999	0.000001	102.875	0.000003	154.005	0.000000
8	114.139	0.000000	109.752	0.000001	104.482	0.000002	98.944	0.000008	149.007	0.000000

c) Message Encrypted by RSA

Table 4.3 Impact Result of Size in RSA of the PSNR value

method	1 LSB		2 LSB		3LSB		4LSB		AFT_LSB	
size	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
1	114.414	0.000000	109.905	0.000001	104.626	0.000002	99.084	0.000008	149.178	0.000000
2	111.511	0.000000	107.248	0.000001	102.340	0.000004	97.024	0.000013	146.714	0.000000
3	109.710	0.000001	105.582	0.000002	100.766	0.000005	95.634	0.000018	145.193	0.000000
8	105.129	0.000002	101.094	0.000005	96.517	0.000015	91.603	0.000045	139.772	0.000000

From tables 4.2 and 4.3 above the encryption is increasing the length of the message and decreasing the value of PSNR. Asymmetric encryption algorithm reduces the value of PSNR more than symmetric see figure 4.2.

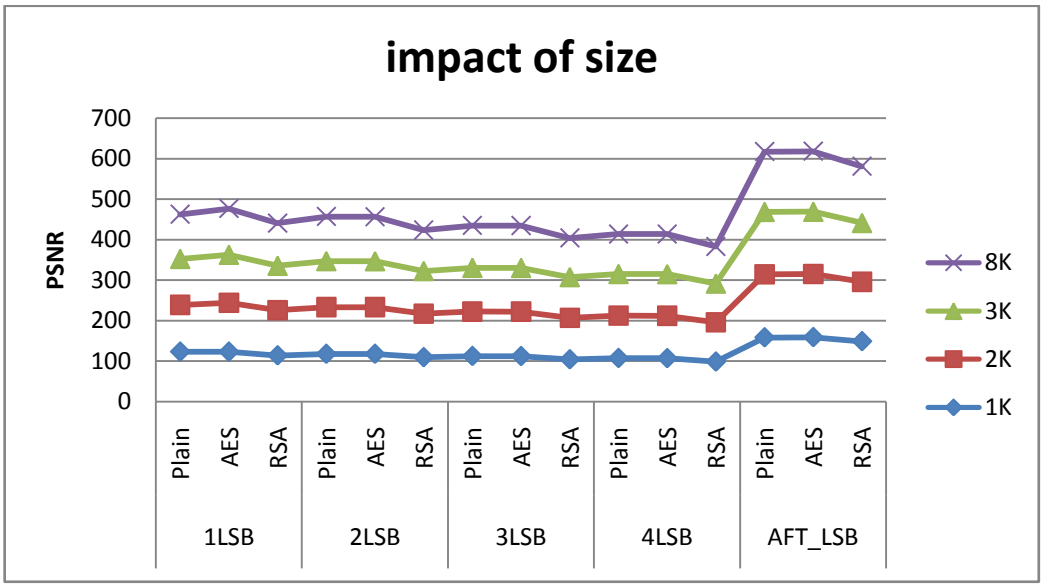


Figure 4.1 Impact of Size

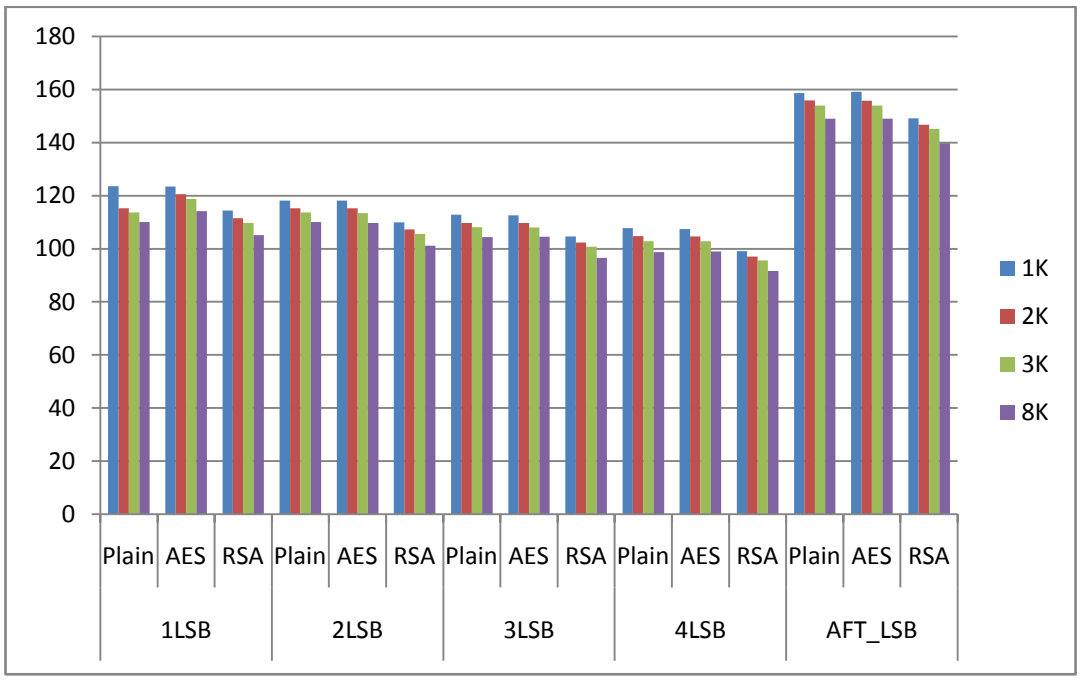


Figure 4.2 Impacts of Encryption

2. Impact of Message File Formats on PSNR value

To study impact of File format I saved 8kb in text, docx and pdf file consecutively and conducted three types of experiments.

a) Message Without Encryption(as plain text)

Table 4.4 impact of message file format on the PSNR value

method	1 LSB		2 LSB		3LSB		4LSB		AFT_LSB	
file format	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
text	124	0.000000	118	0.000000	113	0.000000	108	0.000001	159	0.000000
word	113	0.000000	109	0.000001	104	0.000003	98	0.000010	148	0.000000
pdf	104	0.000003	100	0.000007	95	0.000019	90	0.000059	139	0.000000

b) Message Encrypted by AES

Table 4.5 impact of message file format on the PSNR value

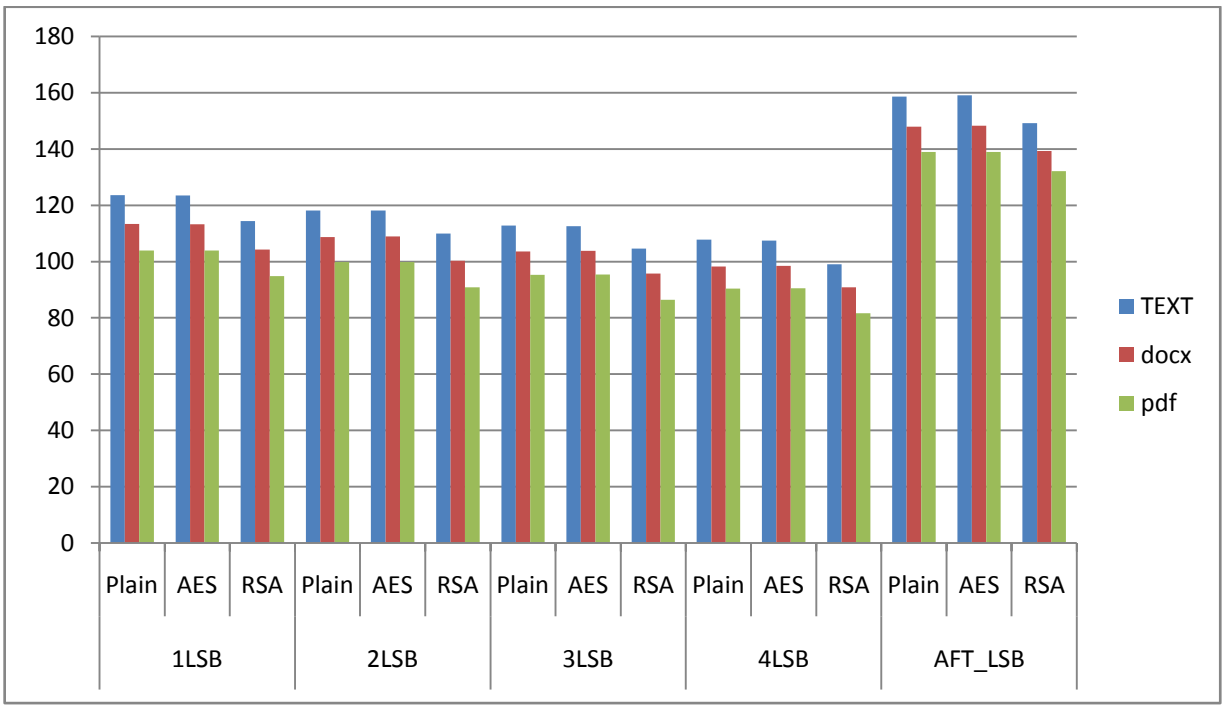
method	1 LSB		2 LSB		3LSB		4LSB		AFT_LSB	
file format	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
text	123	0.000000	118	0.000000	113	0.000000	107	0.000001	159	0.000000
word	113	0.000000	109	0.000001	104	0.000003	98	0.000009	148	0.000000
pdf	104	0.000003	100	0.000007	95	0.000019	90	0.000059	139	0.000000

c) Message Encrypted by RSA

Table 4.6 impact of message file format on the PSNR value

method	1 LSB		2 LSB		3LSB		4LSB		AFT_LSB	
file format	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
text	114	0.000000	110	0.000001	105	0.000002	99	0.000008	149	0.000000
word	104	0.000002	100	0.000006	96	0.000017	91	0.000054	139	0.000000
pdf	95	0.000021	91	0.000053	86	0.000148	82	0.000451	132	0.000000

Tables 4.4, 4.5 and 4.6 above it is clear that the PSNR affected by message file format. In all experiments it is observed the message that saved in PDF format is increasing the message length and decreasing value of PSNR. Figure 4.2 explained impact of message file format.



4.3 Impact of Message File Format

CONCLUSION RECOMMENDATIONS AND FUTURE WORKS

5.1 Conclusion

Steganography technologies are a very important part of the future of Internet security and privacy on open systems.

Hiding digital data such as text messages, documents, and binary files into audio files, such as WAV, is one of the important ways that used to maintain the privacy and confidentiality of data.

One of the methods, that is used to hide data in audio files, is LSB. In this method, the bits of a message are stored in cover consecutively which makes it easy for attackers to extract the hidden message.

In this thesis LSB has been improved by using jumping techniques to store bits of message in cover without affecting the cover size. The new LSB (AFT_LSB) method enhances the value of PSNR more than Standard LSBs. This makes the new method successful in the world of audio Steganography.

Combining Steganography with cryptography added acceptable level of secrecy but it impacts the quality of concealment of the data in ~~the~~ audio files.

5.2 Recommendations

- Messages can be compressed before encrypting it to add an extra level of confidentiality and increase capacity to cover.
- To increase security with less effect on quality we use symmetric Algorithm.
- To decrease the effect of quality on hiding data, we use message saved in text file(.txt)

5.3 Future Works

- This Thesis concentrates only on audio files with ".wav" format. However, it can be extended to a level such that it can be used for the different types of audio wave file formats like .au, .mp3, wma etc.,
- Developing AFT_LSB to use in mobile.
- Conduct experiments on different types of audio like (jaz) to evaluate quality of stegoFile using MSE, PSNR and human hearing.
- To add additional level of security we can encrypt StegoFile key before sending it to the receiver side.

REFERENCES

[1] J.R. Krenn, "Steganography and Steganalysis," January 2004.

[2] Sabu M Thampi-Assistant Professor, "Review, Information Hiding Techniques: A Tutorial," LBS College of Engineering, Kasaragod, Kerala- 671542, S.India,.

[3] Youssef Bassil LACSC – Lebanese Association for Computational Sciences, "A TWO INTERMEDIATES AUDIO STEGANOGRAPHY TECHNIQUE," *Journal of Emerging Trends in Computing and Information Sciences (CIS), ISSN*, vol. 3, no. 11, November 2012.

[4] Dr.V Sundaram, A.Joseph Raphael, "Cryptography and Steganography – A Survey," *ISSN:2229-6093*, vol. 2, pp. 626-630.

[5] Aman Singh Sandeep Singh, "A Review on the Various Recent Steganography," *IJCSN International Journal of Computer Science and Network*, vol. 2, no. 6, December 2013.

[6] Azizah Abdul Manaf2 and Akram M. Zeki3 Abdulaleem Z. Al-Othmani1, "A Survey on Steganography Techniques in Real Time Audio," *IJCSI International Journal of Computer Science Issues*, vol. 9, no. 1, January 2012.

[7] Pratap Chandra Mandal, "Modern Steganographic technique: Asurvey," *Pratap Chandra Mandal / International Journal of Computer Science & Engineering Technology (IJCSET)2229-3345*, vol. 3, no. 9, pp. 444-448, september 2012.

[8] Suresh Gawande2 Rakhi1, "A REVIEW ON STEGANOGRAPHYMETHODS," *International Journal of Advanced Research in Electrical*, vol. 2, no. 10, October 2013.

[9] Ismail Marzuki, Faisal Rahmat Jasril, "CAPACITY ENHANCEMENT OF MESSAGES CONCEALMENTIN IMAGE AND AUDIO STEGANOGRAPHY," *INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEM*, vol. 6, DECEMBER 2013.

[10] K.P.Adhiya Swati A. Patil, "Hiding Text in Audio Using LSB Based Steganography," *ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)*, vol. 2, 2012.

[11] Prof. Samir Kumar Bandyopadhyay1 and Barnali Gupta Banik2, "Multi-Level Steganographic Algorithm forAudio Steganography using LSB Modification and Parity Encoding Technique," *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, vol. 1, no. 2, July – August 2012.

[12] Dimple*, "ENCRYPTION USING DIFFERENT TECHNIQUES: A REVIEW," *International Journal in Multidisciplinary and Academic Research (SSIJMAR)*, vol. 2, January-February 2013.

[13] Nagesh Kumar Jawahar Thakur, "DES, AES and Blowfish: Symmetric Key Cryptography Algorithms Simulation Based Performance Analysis ," *International Journal of Emerging Technology and Advanced EngineeringWebsite: www.ijetae.com (ISSN 2250-2459*, vol. 1, no. 2, December 2011.

[14] Lalit Singh Dr. R.K. Bharti, "Comparative Performance Analysis of Cryptographic Algorithms," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 11, November 2013.

[15] Gregory Kipper, *INVESTIGATOR'S.*

[16] M. Ram Mohan Reddy S.S. Divya, "HIDING TEXT IN AUDIO USING MULTIPLE LSB STEGANOGRAPHY AND PROVIDE SECURITY USING CRYPTOGRAPHY," *INTERNATIONAL JOURNAL OF SCIENTIFIC &*

TECHNOLOGY RESEARCH ISSN 2277-8616, vol. 1, no. 6, JULY 2012.

[17] Venugopala P S Padmashree G, "Audio Stegnography and Cryptography: Using LSB algorithm at 4th and 5th LSB layers," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 2, no. 4, October 2012.

[18] Sumod Tom Philip, Sumaya Nazar, Ashams Mathew & Niya Joseph Shery Elizabeth Thomas, "Advanced Cryptographic Steganography Using Multimedia Files".

[19] Pradeep Kumar Singh Kriti Saroha, "A Variant of LSB Steganography for Hiding Images in," *International Journal of Computer Applications (0975 – 8887)*, vol. 11, December 2010.

[20] (2014, December) [Online]. http://www.ni.com/white-paper/13306/en/

[21] Zhou Wang and Alan C. Bovik. (2014, December) [Online]. https://ece.uwaterloo.ca/~z70wang/publications/SPM09.pdf